The Transition to APR-DRGs and Related Methodological Changes

Health Services Cost Review Commission
4160 Patterson Avenue
Baltimore, MD 21215
(410) 764-2605

June 1, 2005

The purpose of this paper is to propose methodological changes and a transition plan to implement the HSCRC's transition to measuring case mix based on the APR-DRG grouper. This document is a final recommendation and is ready for Commission action.

Executive Summary

The following is a summary of the package of changes proposed by the staff to implement the transition to APR-DRGs:

- Excluded cases: no change in the current logic for excluded cases
- Outlier logic: geometric mean plus 2.5 standard deviations to define outlier cases
    - Cases above the upper limit will be trimmed, and trimmed charges (based on unit rates) will be passed through to the hospital
    - Cases below the lower limit will be excluded and paid on the basis of charges
    - Minimum $10,000 loss in each cell
    - Hospital CPCs will be reconstituted to reflect new outlier rules; changes to the base will be revenue neutral.
- Cases weights: Use hospital relative weights calculated by an iterative approach
- DSH/IME: Regression approach based on individual case mix data
    - DSH measure and IME magnitudes still open for discussion
        - Open questions include how to measure DSH, if the DSH effect differs by urban/rural status or peer groups, if the DSH effect is linear or if the effect changes by level of DSH status.
    - Monitor and refine both DSH and IME over transition the period.
- MDC 19 (psychiatric) cases should be paid differently, probably on a per diem basis.
    - The methodology is still under development.
    - Implementation would be slated for FY07.
- Rehabilitation cases will be referred to a small workgroup for consideration over the Summer, with a recommendation to the Commission in October 2005.
- The proposed governor for ongoing case mix growth for FY06 is
    - <0 100% recognized
    - [0, 1%] 100% recognized
    - (1, 2.5%] 75% recognized
    - (2.5, 4%] 50% recognized
    - >4% 25% recognized
- For future years of a case mix governor, cumulative case mix growth over the entire implementation would be considered to minimize unintended inequities from differences in the timing of case mix growth. Details of the cumulative governor for FY07 and potential future years will be proposed for Commission consideration by the October 2005 meeting.
- Transition plan
    - Moratorium on spenddowns for 4 ROCs (2 years)
    - Moratorium on full reviews for an equivalent period, except temporary reviews in emergency circumstances
    - No scaling during transition
    - Partial rate application process for capital will be referred to a small workgroup for an October 2005 proposal

Introduction

As the Commission moves toward measuring hospital case mix based on the APR-DRG grouper, a number of related methodological issues need to be addressed as part of this transition. The Commission decided to move to APR-DRGs to guarantee that resources in the Maryland hospital system follow patient care, but case mix is only one of the methodologies in the system to measure resource use. Other methodologies such as disproportionate share (DSH) and indirect medical education (IME) are designed to recognize reasonable levels of desirable social costs incurred by a hospital as the institution delivers patient care and fulfills related activities that are part of its overall mission.

The Commission's methodologies, in effect, attempt to measure these costs piece by piece and adjust hospital rates appropriately when comparing a hospital to its peers in the Reasonableness-of-Charges (ROC) Analysis and the Inter-hospital Cost Comparison (ICC) used in a full rate review. Because these adjustments are made sequentially, the results of each adjustment are affected by previous adjustments. For example, because IME is estimated from case-mix adjusted charge-per-case targets, a change in the case mix methodology will affect the amount of IME that should be recognized in hospital rates. To the degree that some costs are reflected in a refined case mix measure, other costs associated with unmeasured patient severity should remain to be explained by the IME adjustment, all else equal.

As the Commission moves to the APR-DRG grouper for measuring case mix, this is the appropriate time to reconsider other methodologies that are related to case mix measurement. This paper presents the staff's recommendation for a number of issues: identifying outlier cases and cases not appropriately reimbursed under a charge-per-case target, constructing case weights for each APR-DRG/severity cell, measuring DSH and IME, handling the transition period to APR-DRGs for the ROC and ICC methodologies, and controlling case mix growth as coding practices improve under this severity-adjusted grouper.

Excluded Cases and Outliers

Traditionally, the HSCRC has excluded some cases from its hospital comparison, passing the costs through to the hospital on a fee-for-service basis. In some cases, these are specialty cases that only one or two institutions in the state serve, such as transplant cases. At other times, unique programs have been excluded because the treatment costs were not adequately covered by the case-mix adjusted payment rates. In other cases, such as burns, the costs were so variable that a methodology based on average payments could not adequately address the high variability associated with the cases in question.

Under APR-DRGs, it is reasonable to reconsider the list of exclusions. Because the grouper is clinically based with four severity levels in each DRG, a reasonable question to ask is whether the logic for specific categories of excluded cases still holds as the Commission shifts to a new case mix methodology.

To assess this question, the case mix workgroup analyzed the dispersion of charges within each APR-DRG severity level. The preliminary analysis suggested that most of the excluded cases could be included without having a substantial effect on the dispersion of charges. The implication is that the logic for excluding the current cases may no longer hold.

In the Case Mix Workgroup meetings, however, some hospital representatives pointed out that transplants offer a great deal of risk to a facility, with many of the associated costs outside the control of the hospital. Prominent examples in transplant cases are the costs of organ acquisition and new technologies that come online. Hospital representatives also noted that the small cell size associated with many transplant cases makes the estimation of weights imprecise and unstable over time. These are reasonable points. Given the small cell size associated with most transplants, the staff recommends that the current list of excluded cases be maintained.[1]

In addition to categorical exclusions, the staff proposes a policy to offer relief for outlier cases. Since the 2000 redesign of the Commission's methodologies, outliers have been defined by the higher of a fixed threshold for charges or a statistical definition for some high-charge cases. In 2000, cases were identified as outliers if they exceeded the outlier threshold which was established on a DRG-by-DRG basis. The threshold was established as the higher of $75,000 or the DRG average charge plus 1.5 standard deviations for the DRG. Any charge above that threshold was treated as a pass-through to the hospital because the case was viewed as medically unmanageable. This approach supplied the hospital with a stop-loss provision on outlier cases to limit the hospital's financial exposure. In the last two rate years, the $75,000 threshold has been inflated by the annual update factor to align the threshold with rising charges statewide.

Under APR-DRGs, this approach has problems. Because Severity 1 cases are generally the least expensive within a DRG and Severity 4 cases are the most severe, a fixed threshold is much more likely to trim substantial revenue from the most severe cases and have little effect on low-severity cases. Consequently, the staff is proposing a revised policy to identify outliers within each APR-DRG/severity cell. Under this proposal, a case would be identified as a high-charge outlier if it exceeded a statistically-determined threshold for each cell, and charges above the threshold would be treated as a pass-through to the hospital. For some cells, the charges show relatively low variability, resulting in a low outlier threshold. Because the intent of this policy is to offer protection against medically unmanageable cases, the statistical definition is modified to guarantee

[1] Kidney transplants have higher volumes and appear much more stable than other transplants, and some members of the case mix workgroup have argued that these cases could be brought under a charge per case target. This paper does not include such a recommendation, however, because the organ acquisition costs are highly variable depending on whether the source of the organ is a cadaver versus a live donor. Differentiating by the source of organ begins to create small cells and unstable weights.

that the minimum loss on any case must be $10,000 before a case is eligible to have charges passed through to the hospital.[2]

Just as the system produces cases that are high-charge outliers, instances of extremely low charges may also be found. These cases are significant in the case mix methodology because, under current methodology, they receive the same credit as all other cases in the DRG/severity cell while consuming fewer resources. In a revenue-neutral system, overpayment to these low-charge cases reduces the revenue available for other cases. To correct this distortion, the staff proposes to exclude these cases from the hospital's CPC and case mix calculation, and the hospital would receive full charges based on HSCRC approved unit rates. Low-charge cases are defined as those that fall more than 2.5 standard deviations below the geometric mean of the DRG/severity cell within the APR-DRG grouper.

For both the new stop-loss thresholds (upper trim limits) and low-charge exclusions, hospital CPC targets must be reconstituted to reflect changes in trim and excluded cases. **This change will be revenue neutral to the hospital as its target is rebased.** A hospital's permanent inpatient revenue at the end of FY05 will be the same as the base inpatient revenue in FY06 before applying the update factor.

Case Weights

A long-standing discussion in the case mix workgroup is how to properly construct case mix weights. Currently HSCRC weights under any grouper are established as follows:

- Remove excluded cases from the data
- Adjust charges for outlier cases so that the maximum charge equals the trim limit
- Calculate an average charge per case in each DRG/severity category
- Calculate a statewide average charge per case
- Divide the cell average by the state average to generate the cell weight

While this method of construction is standard, using aggregate data to construct weights does not necessarily reflect the relative resource use associated with treating patients in any given hospital. If some types of cases are concentrated in high-charge hospitals, the case weights tend to be too high in those case mix cells and weights tend to be too low in less resource intensive cases. This distortion means that weights are too dispersed compared to the actual resource use experienced at most hospitals.

When the case weights are too dispersed, hospitals have an economic incentive to redirect resources toward cases with weights that are too high and to shift resources away from lines of business where case weights are low. This phenomenon occurs because the

[2] The specific statistical methodology is to establish a high-charge threshold for each cell at its geometric mean plus 2.5 standard deviations. If the resulting limit is less than the geometric mean + $10,000, then the upper threshold becomes the higher of the two values.

case weight for a given discharge determines the approved revenue the hospital receives in the Maryland rate-setting system. The approved revenue associated with any case is the hospital's charge-per-case target at a case mix of 1 multiplied by the case weight.[3] All else equal, a higher case weight generates more revenue for the hospital. If weights are too high relative to resource use, these cases will be relatively profitable to the hospital. If weights are too low, cases will be relatively unprofitable. Over time, hospitals have an economic incentive to gravitate toward profitable types of cases and away from losing cases, at least within the hospital's overall mission. Some analysts nationally have suggested that these distortions in weights have contributed to the profitability of cardiac and orthopedic procedures, making these lines of business particularly attractive to hospitals. Conversely, these distortions in weights contribute to the relative unattractiveness of some low-intensity procedures such as obstetrics and psychiatric care.

Conceptually, this issue may be demonstrated by a simple example. Suppose a hospital with relatively high charges has served all the discharges in one DRG/severity cell, and the average charge in the cell is $15,000. Let the hospital's average charge per case be $10,000. The cell weight is 1.5 (cell weight = $15,000/$10,000), meaning that relative resource use in the cell is 1.5 times that of the average case in the hospital. Suppose that the statewide average charge per case is $5,000. Under the current methodology for creating case mix weights, the cell weight is 3, not 1.5 (current weight = $15,000/$5,000).

This illustration demonstrates in stylized fashion a real phenomenon in the construction of case weights. Table 1 below illustrates the difference in case weights based on Johns Hopkins Hospital internal relative weights versus the weights based on the current statewide method. In each of the DRG/severity cells, the current methodology for calculating weights produces a higher weight than would be generated based on the hospital's actual experience.

[3] That is

(CPC/CMI) * case weight = approved revenue associated with the case.

Table 1: Current Weights versus Hospital Internal Relative Weights for Selected DRG/Severity Cells in FY04

hospid	apr_drg	severity	State weight	Hospital internal weight
9	893	3	1.243352	.7967733
9	282	4	2.474957	1.556901
9	443	2	1.325058	.8053566
9	163	3	4.229475	2.582506
9	442	2	1.5796	.9592774
9	21	3	4.211617	2.398296
9	21	4	5.842637	3.417383
9	163	4	6.639062	3.764894
9	480	1	1.100162	.6312057
9	21	1	2.125271	1.245437
9	23	2	1.693878	.9378513
9	21	2	2.65869	1.557991
9	71	2	.7023816	.38283
9	480	2	1.148715	.6536783
9	22	2	1.430847	.8034625
9	303	2	5.246702	3.138875
9	303	3	7.148608	4.103312

The staff has presented analysis to the Case Mix Workgroup for three methods of constructing weights: (1) the current method, (2) a standardized approach, and (3) hospital internal relative weights. While the current method appears to produce weights that are too dispersed, it has formed the basis of the system until now and a number of hospital representatives have questioned the need to modify this methodology, noting its simplicity and the difficulty of identifying any specific methodology as "best". While the staff believes the issue of dispersion needs to be addressed, this approach is the baseline against which other approaches will be measured and has been modeled accordingly.

The second approach, the standardized approach, adjusts hospital charges for markup, labor market differences, disproportionate share, and indirect medical education before calculating case mix weights. By removing some of the variation in charges prior to calculating weights, this approach attempts to reduce the dispersion that results from the concentration of some high charge cases in specific cells. Preliminary analysis demonstrates that this approach does in fact reduce the dispersion in case mix weights and in the resulting case mix indexes for hospitals. However, this approach is unlikely to adjust for all the sources that contribute to differential charges and allows some distortion to remain in the case weights. To the degree that the adjustments are incorrect, the standardization approach could inadvertently introduce distortions of its own.

Some have argued that indirect medical education should not be removed as part of the standardization process, contending that it is part of the cost of treating the patient. In Maryland and other payment systems, however, IME is treated as an aggregate estimated amount that is recognized across all the hospital's rates. The hospital-level cost is spread across all rates as a markup over patient-service costs, and is not measured at the specific patient or revenue-center level. In Maryland, this single factor is the source

of most of the distortion in weights. If IME is not removed, standardized weights are similar to the current weights. If IME is removed, standardized weights are similar to the hospital relative weights (discussed below).

The third approach – the hospital relatives approach – attempts to directly address the remaining distortion. This approach is designed to calculate case weights based on hospital's own internal relative weights. The methods of computation are complex and are not described in detail in this paper. Conceptually, however, the intent of these methods is to compute weights based on the internal distribution of hospital relative resource use. In the Case Mix Workgroup, these methods have been referred to as the "iterative approach" because an initial set of weights is constructed and is then modified until a stable set of weights is calculated.

A reasonable question put forth in the workgroup is how do we know which set of weights is best. Each method produces case weights that can be used to calculate a case mix index for each hospital. Under each approach some hospitals fare better and others fare worse. What are the criteria for choosing which method best approximates resource use in Maryland hospitals?

The staff has attempted to address this issue empirically by measuring which of the three approaches produces weights that best fit the internal weights observed from each Maryland hospital. To do so, the staff calculated APR-DRG/severity weights from Maryland FY04 data based on the data from each hospital. For each cell, the staff calculated the difference between the cell weight from state data based on each weighting methodology and the cell weight from the hospital's own data, weighting for the number of cases in each hospital cell. After taking the absolute value of each difference, these differences (or deviations) were added up for every cell across all hospitals.[4] This approach was designed to determine which methodology produces weights that best reflect the internal resource use within hospitals. The results for FY04 show that the hospital relative approach best fit the internal weights of hospitals overall. The current methodology had the worst fit. The standardized approach fell between these two. These results support the case that the hospital relative approach is superior for calculating case mix weights. See Appendix I for a summary of these results.

Other case mix issues relate to psychiatric and rehabilitation cases. The APR-DRG grouper does not explain the variation in charges for psychiatric cases any better than the current methodology, and generally psychiatric cases in other states have not been paid on a per case basis because of the variability in their length of stay. The risk associated with the cases results in limited access to care and financial risk for the institutions who serve these patients. Both Johns Hopkins Hospital and CareFirst have suggested that these cases be paid on a per diem basis instead on a per case basis as is current policy. The staff concurs that these cases are not handled well under the current methodology and proposes that cases in MDC-19 be treated under a separate methodology, likely a per diem system. A number of operational issues exist before this change can be accomplished, and we believe that we will not be able to complete an

[4] The measure was also calculated using squared deviations with qualitatively similar results.

alternative methodology for FY06. As part of a package to more properly measure and pay for system resources, however, a commitment to improve this methodology by FY07 is appropriate at this time.

For rehabilitation cases, a few hospitals have noted that APR-DRGs are not as sensitive a measure as the existing HSCRC logic for handling these cases. Because this issue has not received sufficient attention at this time, the staff will address this issue over the Summer and propose any necessary changes by the October 2005 Commission meeting.

Disproportionate Share and Indirect Medical Education

Discussions between staff, hospitals, and payers about disproportionate share (DSH) and indirect medical education (IME) have been ongoing for some time. A number of hospital representatives have argued that the DSH adjustment is too small, but analyses of DSH costs generally showed small and often statistically insignificant DSH effects. Additionally, there has been much dispute about the estimation and calculation of the IME adjustment. Determining appropriate magnitudes for either of these adjustments is difficult, particularly because separating the DSH and IME effects is difficult. Frequently, hospitals with high DSH measures have substantial teaching as well, so separating DSH and IME related costs is a difficult statistical problem.

The model proposed by the staff combines a disproportionate share and indirect medical education adjustment that is estimated from the hospital discharge data. The model introduces a number of control variables for patient age, gender, source of admission, patient disposition, and payer while including hospital-level measures for DSH and IME. The exact form of the measure for DSH has not been resolved, but the most promising candidate is the percent of the hospital's discharges that are classified as Medicaid, Medicare with Medicaid as a secondary payer, and self-pay patients. Other forms based on the income levels of zip codes from which a hospital's patients are drawn is also under consideration. The DSH and IME results are estimated by MDC, and the overall result is combined for a single DSH/IME adjustment for each hospital.

The preliminary results of this analysis appear promising. Regression results historically have been unable to find a DSH effect while the IME adjustment has been large. This methodology allocates more toward DSH and less toward IME. The net effect for teaching hospitals still results in a large adjustment because teaching hospitals often have large DSH populations. However, this approach allows non-teaching hospitals with large poor populations larger adjustments for DSH while the current methodology attributes almost all the credit toward teaching. The final result in the ROC analysis, however, depends on the hospital's adjustment versus the rest of the peer group.

Some community hospitals have noted their concern about the use of regression methods for calculating DSH and IME. Because various DSH measures demonstrate sensitivity, some hospital representatives asked for further analysis before settling on a specific measure or approach. Others were concerned about the dynamic implications of

IME. Because IME is measured as the residual variation in charges correlated with a teaching measure (the ratio of approved residents per occupied bed, or RESBED), higher charges for any reason at teaching hospitals (particularly academic medical centers who have high RESBED ratios) may translate into higher IME measurements over time.

This concern is legitimate. While staff believes this regression approach modeled for the Case Mix Workgroup is a substantial improvement over the current methodology, it is subject to the same bias over time as the simple regression methodology currently in use. The staff proposes that this model based on the case mix data be used to measure DSH and IME, but the results for both measures should be monitored during the phase-in period for APR-DRGs. Further refinements in the DSH measure would be explored during this period. Some questions to consider are the form of the DSH measure, whether the effect differs for urban and rural facilities or by peer group, and whether DSH should be modeled as a nonlinear (or perhaps non-continuous) relationship with hospital costs. Additionally, the level of IME would be reported as part of the ROC to explicitly show the amount of revenue, revenue per approved resident, and revenue per discharge that is associated with the IME adjustment. Further refinements could then be discussed as necessary if IME growth diverged excessively from overall revenue growth in the state. These statistics will provide full disclosure for the system instead of burying the IME adjustment in a formulaic calculation.

Case Mix Governors

Another issue that has received much attention is how to structure a governor on case mix growth in the transition to APR-DRGs. Because we expect case mix growth to rise rapidly as hospitals code more completely under the new grouper, the staff believes that controls on case mix growth must be established to keep the system within its overall goals for revenue growth. After discussing possible options with the Case Mix Workgroup, the staff has developed a recommendation designed to deliver full case mix growth at low levels and recognize only portions of case mix growth above certain thresholds. Because the update factor for all hospitals reflects expected case mix growth of 1.7% for FY06, we have attempted to develop a governor on case mix growth to remain within that budget. This model was calibrated by looking at annualized rates of case mix growth under APR-DRGs in the current fiscal year. During this time period, statewide case mix growth is 2.8%. This governor would deliver 1.5% growth. We deliberately chose a conservative measure because we expect coding practices to become more aggressive when hospital revenue will be directly affected by coding improvements. The governor is structured as follows for FY06. In future years, this structure may need to be recalibrated based on actual experience and future revenue growth expectations.

- Governor for ongoing case mix growth
  - <0 100% recognized
  - [0, 1%] 100% recognized
  - (1, 2.5%] 75% recognized
  - (2.5, 4%] 50% recognized
  - >4% 25% recognized

Because some hospitals may be restricted from receiving case mix growth that is related to programmatic changes, not coding improvements, a portion of the update factor has been set aside for FY06 to restore some of this case mix growth. Hospitals opening new programs or expanding resource-intensive services could see substantial case mix increases even under the Maryland version of the CMS grouper. To recognize that case mix growth of this sort is different than the growth associated with coding improve, the staff would assess programmatic requests on a case-by-case basis.

For future years of a case mix governor, cumulative case mix growth over the entire implementation would be considered to minimize unintended inequities from differences in the timing of case mix growth. Details of the cumulative governor for FY07 and potential future years will be proposed for Commission consideration by the October 2005 meeting.

Phase In

A number of possibilities for phasing in these methodologies exist, and there is general agreement in the Case Mix Workgroup that volatility resulting from these methodology changes should be minimized as these changes are being introduced to the system. The staff believes that the most viable option for introducing these changes is to introduce a moratorium on spend downs, full reviews, and scaling during this transition period. Because coding improvements will likely take some time, we believe an appropriate period of time would cover four ROCs or two fiscal years of data. This solution seems to introduce the fewest problems during implementation, although problems still remain. Issues also remain around hospitals' access to the partial rate methodology for capital during the moratorium.

The staff recommends that the moratorium for full reviews, spenddowns, and scaling become effective November 1, 2005 and end November 1, 2007. As for partial rate reviews for capital, the staff recommends that this issue also be addressed by a workgroup over the Summer for a recommendation at the October 2005 Commission meeting. While the staff is skeptical that a credible approach can be developed to establish reasonable rates within the partial rate process during the moratorium, the MHA has noted that its membership has been unable to address this issue due to the extended deliberations around other methodological issues. Considering partials as part of the abbreviated Summer discussion will provide the hospital representatives the opportunity to address this issue.

Appendix I: RELATIVE WEIGHT DIFFERENCES BETWEEN HOSPITAL INTERNAL WEIGHTS AND OTHER METHODS OF WEIGHT CALCULATION (USING FISCAL YEAR 2004 DATA)

Method	Absolute Deviation	Squared Deviation	% Greater than Iterative: Absolute Deviation	% Greater than Iterative: Squared Deviation
Current	167,802	190,211	7.24%	14.36%
Standardized	159,246	167,839	1.78%	0.91%
Hospital Relative (iterative)	156,468	166,329	0.00%	0.00%

Note: The best fit in this analysis has the lowest total deviations.